

TITLE OF THE INVENTION

Semiconductor Memory System Having Multiple System Data Buses

BACKGROUND OF THE INVENTION

5 1. Field of the Invention

[0001] The present invention generally relates to the field of semiconductor memory systems, and more particularly, the present invention relates to a semiconductor memory system having multiple system data buses.

10 [0002] A claim of priority is made to Korean Patent Application No. 2002-50172, filed on August 23, 2002, in the Korean Intellectual Property Office, the disclosure of which is incorporated herein in its entirety by reference.

2. Description of the Related Art

15 [0003] Memory devices have been developed with emphasis on the realization of high degrees of integration and large memory capacities. Central processing units (CPUs), on the other hand, have been developed with emphasis on achieving high processing speeds. As a result, the difference in operating speeds between CPUs and memory devices has been gradually increasing. As such, the operating speed of the computer system memory device has become a chief factor limiting the performance of the computer system itself.

20 [0004] FIG. 1 is a block diagram of the structure of a conventional semiconductor memory system.

25 [0005] A system data bus having a width of M bits connects memory modules 120, 130, 140, and 150 to a memory controller 110. The memory modules 120, 130, 140, and 150 are installed in respective memory slots (not shown). The memory modules 120, 130, 140, and 150 have the same data bus width as the system data bus. Each of the memory slots is commonly connected to the system data bus.

30 [0006] Also, all of the memory slots are sequentially connected to the system data bus. Thus, in large-capacity memory systems, as the number of memory slots increases, channel discontinuity in the system data bus line and impedance mismatch increase. Hence, the operating characteristics of the system

data bus are deteriorated in a high frequency region, thus limiting the operation speed of the system.

[0007] Large-capacity memory devices are essential for, in particular, the operation of a high-performance, large-capacity server system. In order to 5 maximize the capacity of a memory system, the number of memory devices installed in the memory system should be maximized. For this purpose, the number of memory slots accommodating memory modules should be increased, and the number of memory devices per memory module should be maximized.

[0008] On the other hand, for high-speed operation of the memory system, 10 an impedance mismatch factor of the system data bus or load of the system data bus caused by the memory devices should be minimized. In order to meet these requirements for high-speed operation, both the number of memory slots connected to a channel and the number of memory modules should be small.

[0009] Thus, a conflict exists between maximizing the number of memory 15 devices to achieve a large memory capacity and minimizing the number of memory devices to achieve a high-speed operation. It is therefore difficult to simultaneously achieve both a large memory capacity and a high-speed operation.

SUMMARY OF THE INVENTION

[0010] According to an aspect of the present invention, there is provided a 20 semiconductor memory system. The semiconductor memory system includes a memory controller, N system data buses (where N is a natural number of 2 or more), and first through P -th memory module groups (where P is a natural number).

[0011] The N system data buses are connected to the memory controller, 25 and each of the N system data buses has a width of M/N bits (where M is a natural number). The first through P -th memory module groups are connected to the N system data buses, and each of the first through P -th memory module groups has N memory modules.

[0012] In each of the first through P -th memory module groups, a different 30 one of the N system data buses is connected to each of the N memory modules.

[0013] The first through P -th memory module groups are operated in response to first through P -th corresponding chip select signals.

[0014] M is the bit-width of an entire system data bus of the semiconductor memory system. The N system data buses are wired such that data transmission times are the same from each N memory modules that operate in response the same chip select signal to the memory controller.

5 [0015] L memory devices may be installed in each of the memory modules, with each of the L memory devices has a bus width of $M/(N*L)$ bits.

10 [0016] Alternately, the L memory devices installed in each of the memory modules may be divided into N banks, with each of the L memory devices having a bus width of M/L bits. Each of the N banks shares a data bus of M/N bits and is separately operated in response to the first through P -th corresponding chip select signals.

15 [0017] According to another aspect of the present invention, there is provided a semiconductor memory system. The semiconductor memory system includes a memory controller, N system data buses (where N is a natural number of 2 or more), a first memory module group, and a second memory module group.

20 [0018] The N system data buses are connected to the memory controller, and each of the N system data buses has a width of M/N bits (where M is a natural number). The first memory module group is connected to the N system data buses, having N memory modules. The second memory module group is connected to the N system data buses, having one or more memory modules.

[0019] In the first memory module group, a different one of the N system data buses is connected to each memory module, has a data bus width of M/N bits and is operated in response to the same chip select signal.

25 [0020] In the second memory module group, all of the N system data buses are connected to each memory module, each of the N system data buses has a data bus width of M bits, and each memory module is operated in response to a separate chip select signal.

30 [0021] M is the bit-width of an entire system data bus of the semiconductor memory system. The N system data buses are wired such that data transmission times are the same from each N memory modules that operate in response the same chip select signal to the memory controller.

[0022] L memory devices may be installed in each of the memory modules, with each of the L memory devices has a bus width of $M/(N*L)$ bits.

[0023] Alternately, the L memory devices installed in each of the memory modules may be divided into N banks, with each of the L memory devices having a bus width of M/L bits. Each of the N banks shares a data bus of M/N bits and is separately operated in response to the first through P -th corresponding chip select signals.

[0024] According to another aspect of the present invention, there is provided a semiconductor memory system which employs a serial link bus system. The system includes a memory controller, N system data buses, and first through P -th memory module groups.

[0025] The N system data buses are connected to the memory controller, and each of the N system data buses has a width of M/N bits. The first through P -th memory module groups are connected to the N system data buses, and each of the first through P -th memory module groups has N memory modules. In each of the first through P -th memory module groups, a different one of the N system data buses is connected to each of the N memory modules, and each of the N system data buses has a data bus width of M/N bits and includes data buffers for data transmission. The first through P -th memory module groups are operated in response to first through P -th corresponding chip select signals, and each of the memory modules connected to a common system data bus transmits data using the data buffers.

[0026] M is the bit-width of an entire system data bus of the semiconductor memory system. The N system data buses are wired such that data transmission times are the same from N memory modules that operate in response the same chip select signal to the memory controller.

[0027] L memory devices may be installed in each of the memory modules, with each of the L memory devices has a bus width of $M/(N*L)$ bits.

[0028] Alternately, the L memory devices installed in each of the memory modules may be divided into N banks, with each of the L memory devices having a bus width of M/L bits. Each of the N banks shares a data bus of M/N bits and is

separately operated in response to the first through P -th corresponding chip select signals.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029] The above and other aspects and advantages of the present invention will become more apparent from the detailed description that follows, with reference to the accompanying drawings, in which:

[0030] FIG. 1 is a block diagram of the structure of a conventional semiconductor memory system;

[0031] FIGS. 2A through 2D are block diagrams of semiconductor memory systems according to a first embodiment of the present invention;

[0032] FIGS. 3A and 3B illustrate alternative memory modules which may be used in the systems of FIGS. 2A through 2D;

[0033] FIGS. 4A and 4B are block diagrams of semiconductor memory systems according to a second embodiment of the present invention;

[0034] FIG. 5A illustrates a conventional semiconductor memory system employing a serial link bus method; and

[0035] FIG. 5B illustrates a semiconductor memory system employing a serial link bus system according to a third embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0036] Hereinafter, preferred embodiments of the present invention will be described in detail with reference to the accompanying drawings. Like reference numerals refer to like elements throughout the drawings.

[0037] FIGS. 2A through 2D are block diagrams of semiconductor memory systems according to a first embodiment of the present invention. FIGS. 3A and 3B illustrate alternative memory modules which may be used in the systems of FIGS. 2A through 2D.

[0038] Referring to FIG. 2A, a semiconductor memory system 200 according to the first embodiment of the present invention includes a memory controller 201, N system data buses 215 and 217, and first through P -th memory module groups 203 and 209.

[0039] An M -bit bus system is defined by N system data buses 215 and 217 connected to the memory controller 201, where each of the N system data buses 215 and 217 has a width of M/N bits. Here, M is the width of an entire system data bus of the semiconductor memory system 200. The N system data buses 215 and 217 are wired such that data transmission times to the memory controller 201 from those of N memory modules 205, 207, 211, and 213 that operate in response to the same chip select signal CS1 or CS2 are the same. In this example, N is 2. That is, there are two system data buses 215 and 217.

[0040] The first through P -th memory module groups 203 and 209 are connected to the system data buses 215 and 217. Further, each group 203 and 209 includes N memory modules 205 and 207, and 211 and 213, respectively. Within each of the first through P -th memory module groups 203 and 209, each of the memory modules 205 and 207, and 211 and 213 is connected to a different one of the system data buses 215 and 217. In this example, P is 2. That is, there are first and second memory module groups 203 and 209. The first through P -th memory module groups 203 and 209 are operated in response to first through P -th chip select signals CS1 and CS2.

[0041] The memory modules 205, 207, 211, and 213 of FIG. 2A may be configured as shown in FIG. 3A or as shown in FIG. 3B.

[0042] In the case of FIG. 3A, L memory devices are installed in the memory modules 300, with each of the L memory devices having a data bus width of $M/(N*L)$ bits. The L memory devices are operated in response to a chip selection signal CS1.

[0043] Alternatively, referring to FIG. 3B, the L memory devices installed in the memory module 310 are divided into N banks B1 and B2, with each of the L memory devices having a data bus width of M/L bits. Each of the N banks B1 and B2 shares a data bus of M/N bits and is separately operated in response to a corresponding one of chip select signals CS1_A and CS1_B. The chip select signals CS1_A and CS1_B are respectively applied to the N banks B1 and B2.

[0044] The N system data buses 215 and 217 of the semiconductor memory system 200 of FIG. 2A are connected to the memory controller 201, and each of the N system data buses 215 and 217 has a width of M/N bits. In this example, N

is 2, and therefore two system data buses 215 and 217 are shown in FIG. 2A.

However, N may be more than 2.

[0045] Also, M is the bit-width of an entire system data bus of the semiconductor memory system 200. Since the width of the entire system data bus is M bits, each of the system data buses 215 and 217 has a width of $M/2$ bits. For example, in the case where the semiconductor memory system 200 includes three system data buses, each of the system data buses has a data width of $M/3$ bits.

[0046] The first through P -th memory module groups 203 and 209 are connected to the system data buses 215 and 217. The memory module group 203 includes the N memory modules 205 and 207, and the memory module group 209 includes the N memory modules 211 and 213. In the first embodiment, P is 2, and accordingly, there are the first and second memory module groups 203 and 209. However, P is not limited to 2.

[0047] In the first memory module group 203, a different one of the system data buses 215 and 217 is connected to each of the memory modules 205 and 207, and in the second memory module group 209, a different one of the system data buses 215 and 217 is connected to each of the memory modules 211 and 213. As stated previously, each of the system data buses 215 and 217 has a data bus width of M/N bits. Therefore, each of the memory modules 205, 207, 211, and 213 has a data bus width of $M/2$ bits.

[0048] The entire system data bus of M bits cannot be filled with one memory module, and thus is filled with N memory modules. In the present embodiment, N is 2 and thus, the entire system data bus is filled with two memory modules. The number of memory modules which can fill M bits constitutes the number of memory modules in a memory module group. Here, two memory modules can fill the width of the system data bus of M bits, and thus, each memory module group has two memory modules.

[0049] Memory modules belonging to the same memory module group are simultaneously operated in response to the same chip select signal. In this embodiment, the first through P -th memory module groups 203 and 209 are operated in response to the first through P -th chip select signals CS1 and CS2, respectively.

[0050] The N system data buses 215 and 217 are wired such that data transmission times are the same to the memory controller 201 from those of N memory modules 205, 207, 211, and 213 that operate in response the same chip select signal CS1 or CS2. That is, since the memory module 205 is closer to the memory controller 201 than the memory module 207, a data transmission time from the memory module 205 to the memory controller 201 may be different than that from the memory module 207 to the memory controller 201. However, by deliberately increasing the length of the wiring of the system data bus 215 from the memory module 205 to the memory controller 201, data transmission times from the two memory modules 205 and 207 to the memory controller 201 can be made the same.

[0051] The semiconductor memory system 220 of FIG. 2B employs another arrangement of memory modules. In the semiconductor memory system 220 of FIG. 2B, one memory module group 223 is comprised of four memory modules 225, 227, 229, and 231. That is, P is 1, and N is 4. Each of system data buses 233, 235, 237, and 239 has a width of $M/4$ bits. The memory modules 225, 227, 229, and 231 are simultaneously operated in response to the same chip select signal CS1. If the system data buses 233, 235, 237, and 239 are wired to the same lengths, the respective data transmission speeds from the memory modules 225, 227, 229, and 231 to the memory controller 221 can be made the same.

[0052] If system data buses from memory modules to a memory controller are wired to the same length, the memory modules may be arranged in a variety of different ways. As an example, in FIG. 2C, the first of two data buses is connected to the first two memory modules, and the second of the two data buses is connected to the second two memory modules. As another example, in FIG. 2D, the first of two data buses is connect to the first and fourth memory module, and the second of the two data buses is connected to the third and fourth memory modules.

[0053] According to the first embodiment of the present invention, by dividing the system memory bus into multiple data buses each connected a memory module contained in each memory module group, memory channel

discontinuity and impedance mismatch are greatly reduced compared with the conventional semiconductor memory system 100 shown in FIG. 1.

[0054] In addition, the high-speed characteristics of the system data bus are improved, and the system data bus can operate in a high-frequency region.

5 Further, the number of data input/output circuits that simultaneously operate in each memory module is reduced to half (or more) of their original number, such that operating power consumption per memory module is reduced. Also, radio frequency noise such as simultaneous switching noise can be reduced.

10 [0055] Furthermore, the number of data pins of the memory modules is reduced such that the number of pins of a memory module socket is reduced, or remaining pins may be used only to ground the memory modules and for power pins, thereby improving the characteristics of the memory modules.

15 [0056] Data bus widths of the memory modules 205, 207, 211, and 213 used in the first embodiment of the present invention shown in FIG. 2 can be reduced to half that of the memory modules 120, 130, 140, and 150 of the conventional semiconductor memory system 100 of FIG. 1.

20 [0057] For this purpose, reducing the number of memory devices of the conventional memory modules 120, 130, 140, and 150 to half of their original number is the simplest method. However, this method is not preferable because it reduces the capacity of a memory module and, consequently, of the entire semiconductor memory system.

[0058] In order to avoid these problems, one data pin is shared by two memory devices in the memory modules, and the two memory devices are operated as separate memory banks.

25 [0059] That is, the data bus width of the memory modules according to the present invention is reduced to half of the data bus width of the conventional memory modules, but the number of memory banks is increased to twice the number of memory banks of the conventional art. In this way, memory capacity can be maintained the same as in the conventional art, and simultaneously, the above-described advantages of the system data bus according to the present invention shown in FIG. 2A can be realized.

[0060] The memory module 300 of FIG. 3A is a conventional memory module. The memory module 300 includes eighteen memory devices. For example, if the memory module 300 has a 72-bit data bus, each of the memory devices has a 4-bit data bus. The memory devices are simultaneously operated in response to the chip select signal CS1.

[0061] The memory module 310 of FIG. 3B is a memory module according to the present embodiment. L memory devices are installed in the memory module 310. The L memory devices are divided into N banks B1 and B2, and each of the N banks B1 and B2 has L/N memory devices.

[0062] The memory module 310 preferably has a 36-bit data bus. Similarly, the memory module 310 includes eighteen memory devices, like the conventional memory module 300. That is, L is 18. Here, N is 2. Thus, the memory module 310 includes two memory banks B1 and B2, and each of the two memory banks B1 and B2 has nine memory devices.

[0063] Each of the two memory banks B1 and B2 shares a 36-bit data bus of the memory module 310 and is separately operated in response to a corresponding one of chip select signals CS1_A and CS1_B. The chip select signal CS1_A is applied to the bank B1, and the chip select signal CS1_B is applied to the bank B2. Each of the eighteen memory devices has a data bus width of 4 bits.

[0064] If the bank B1 is selected in response to the chip select signal CS1_A, the nine memory devices of the bank B1 transmit data to a memory controller using the 36-bit data bus. In this case, the bank B2 is not operated. If the bank B2 is selected in response to the chip select signal CS1_B, the nine memory devices of the bank B2 transmit data to the memory controller using the 36-bit data bus. In this case, the bank B1 is not operated.

[0065] In this way, memory capacity can be maintained the same as in the conventional semiconductor memory system, and simultaneously, the advantages of the present invention described previously can be realized.

[0066] If there are four system data buses like in the memory module 220 of FIG. 2B, that is, if N is 4, the number of memory devices of the memory module 220 is the same as in the conventional art, while the number of memory banks is

increased four times. Memory devices belonging to four different memory banks share a data pin of a memory module. In this way, the data bus width is reduced to 1/4, and simultaneously memory capacity can be maintained the same as in the conventional art.

5 [0067] Another method for reducing the data bus width of the memory module without reducing memory capacity is to include in memory module memory devices having a narrow data bus width. If memory devices having a data bus width half of an original data bus width are used, the data bus width of a memory module having the same number of memory devices is also reduced by half.

10 [0068] That is, the memory module 200 of FIG. 2A includes eighteen memory devices, and each of the eighteen memory devices has a data bus width of $M/(N*L)$ bits, i.e., $72/(2*18) = 2$ bits.

15 [0069] In this way, a memory module whose data bus width is reduced to half of its original width and which has the same memory capacity and the same number of memory banks can be constituted.

20 [0070] FIGS. 4A and 4B are block diagrams of a semiconductor memory system according to a second embodiment of the present invention. Referring to FIG. 4A, a semiconductor memory system 400 according to the second embodiment of the present invention includes a memory controller 410, N system data buses 423, 425, and 427, and first and second memory module groups 413 and 420.

25 [0071] The N system data buses 423, 425, and 427 are connected to the memory controller 410, and each of the N system data buses 423, 425, and 427 has a width of M/N bits. Here, M is the bit-width of an entire system data bus of the semiconductor memory system 400. The N system data buses 423, 425, and 427 are wired such that data transmission times from memory modules 415, 417, and 419 of the first memory module group 413 to the memory controller 410 are the same.

30 [0072] The first memory module group 413 is connected to the N system data buses 423, 425, and 427 and includes the N memory modules 415, 417, and 419. The second memory module group 420 is connected to the N system data buses 423, 425, and 427 and includes one or more memory modules 421.

[0073] In the first memory module group 413, a different one of the system data buses 423, 425, and 427 is connected to each of the memory modules 415, 417, and 419, and each of the system data buses 423, 425, and 427 has a data bus width of M/N bits and is operated in response to the same chip select signal CS1.

[0074] More specifically, L memory devices (not shown) are installed in the memory modules 415, 417, and 419 of the first memory module group 413. The L memory devices are divided into N banks (not shown), and each of the N banks has L/N memory devices. Each of the N banks shares a data bus of M/N bits of the memory modules 415, 417, and 419 and is separately operated in response to a corresponding chip select signal.

[0075] Each of the L memory devices has a data bus width of M/L bits.

[0076] In the memory module 421 of the second memory module group 420, the N system data buses 423, 425, and 427 are connected to the memory module 421, and each of the system data buses 423, 425, and 427 has a data bus width of M bits, and each memory module is operated in response to a separate chip select signal CS2.

[0077] L memory devices are installed in the memory modules 415, 417, and 419 of the first memory module group 413. Each of the L memory devices has a data bus width of $M/(N*L)$ bits.

[0078] Hereinafter, the operation of the semiconductor memory system according to the second embodiment of the present invention will be described in detail with reference to FIGS. 4A and 4B.

[0079] The semiconductor memory systems 400 and 430 of FIG. 4A and 4B according to the second embodiment of the present invention represent a case where memory modules having different data bus widths are used together.

[0080] Referring to FIG. 4A, the entire system data bus having a width of M bits is divided into the three system data buses 423, 425, and 427 having a width of $M/3$ bits, and there are the three memory modules 415, 417, and 419 in the first memory module group 413. Each of the three memory modules 415, 417, and 419 has a data bus width of $M/3$ bits. The structure and operation of the first memory

module group 413 are the same as those of the semiconductor memory system according to the first embodiment of the present invention.

[0081] The second memory module group 420 is connected to the three system data buses 423, 425, and 427 and includes a memory module 421.

5 However, the second memory module group 420 may include two memory modules 441 and 443, like a second memory module group 440 of FIG. 4B, or more than two memory modules.

[0082] The memory module 421 of the second memory module group 420 has the same data bus width of M bits as the entire system data bus and is 10 operated in response to the chip select signal CS2.

[0083] The memory modules 415, 417, and 419 of the first memory module group 413 have the same data bus widths as the divided system data buses 423, 425, and 427, and the memory module 421 of the second memory module group 420 has the same data bus width as the data bus width of the entire system data 15 bus.

[0084] Referring to FIG. 4B, a first memory module group 433 includes two memory modules 435 and 437 and is connected to two divided system data buses 450 and 453. The second memory module group 440 includes two memory modules 441 and 443. Each of the memory modules 441 and 443 has a data bus width of M bits and is connected to the two divided system data buses 450 and 453. 20

[0085] Like the semiconductor memory systems 400 and 430 according to the second embodiment of the present invention, memory modules having different data bus widths can be combined with one another in various ways, and the same effect as that of the first embodiment can be obtained.

25 [0086] FIG. 5A illustrates a conventional semiconductor memory system employing a serial link bus. FIG. 5B illustrates a semiconductor memory system employing a serial link bus according to a third embodiment of the present invention.

[0087] Referring first to FIG. 5B, a semiconductor memory system 560 employing a serial link bus according to the third embodiment of the present 30 invention includes a memory controller 565, N system data buses 567 and 569, and first through P -th memory module groups 570 and 575.

[0088] The N system data buses 567 and 569 are connected to the memory controller 565, and each of the N system data buses 567 and 569 has the width of M/N bits. Here, M is the width of an entire system data bus of the semiconductor memory system 560. The N system data buses 567 and 569 are wired such that data transmission times are the same from the memory controller 565 to the N memory modules 571, 573, 577, and 579 that operate in response to the same chip select signals CS1 and CS2.

[0089] The first through P -th memory module groups 570 and 575 are connected to the N system data buses 567 and 569, and each of the first through P -th memory module groups 570 and 575 includes the N memory modules 571 and 573, and 577 and 579, respectively. In each of the first through P -th memory module groups 570 and 575, a different one of the system data buses 567 and 569 is connected to each of the memory modules 571, 573, 577, and 579, and each of the system data buses 567 and 569 has a data bus width of M/N bits and includes data buffers DBN1, DBN2, DBN3, and DBN4 for data transmission.

[0090] The first through P -th memory module groups 570 and 575 are operated in response to corresponding ones of the first through P -th chip select signals CS1 and CS2, and each of the memory modules 571, 573, 577, and 579 connected to the common system data bus transmits data using the data buffers DBN1, DBN2, DBN3, and DBN4.

[0091] The L memory devices (not shown) are installed in the N memory modules 571, 573, 577, and 579, with each of the L memory devices having a data bus width of $M/(N*L)$ bits.

[0092] Alternately, the L memory devices installed in the N memory modules 571, 573, 577, and 579 may be divided into N banks (not shown), each of the L memory devices having a data bus width of M/L bits. Each of the N banks shares a data bus of M/N bits of the N memory modules 571, 573, 577, and 579 and is separately operated in response to the corresponding chip select signals CS1 and CS2.

[0093] Hereinafter, the operation of the semiconductor memory device according to the third embodiment of the present invention will be described in detail with reference to FIGS. 5A and 5B.

[0094] The conventional serial link bus of FIG. 5A uses a data transmission technique in which signals are sequentially transmitted from a previous memory module to a next memory module using serial link when memory modules 520, 530, 540, and 550 of the semiconductor memory system 500 are connected to a system data bus 555.

[0095] A memory module receives signals from the previous memory module and transmits the signals to the next memory module at a next clock cycle. Such data transmission may be directly performed between memory devices or through an additional data buffer for data transmission.

[0096] Since signal transmission in the serial link bus is performed point-to-point between two adjacent memory devices or data buffers, signal fidelity is very high even in high-speed operation.

[0097] However, each time data is transmitted to the next memory module, at least one or more clock cycles of time is required for a data buffering operation. Thus, the latency of the semiconductor memory system is increased, reducing the operating speed of the semiconductor memory system.

[0098] In particular, in a large-capacity semiconductor memory system, as the number of memory modules is increased, the number of intermediate buffering operations performed is increased. Thus, time delay caused by the intermediate buffering periods is increased, and the speed of the semiconductor memory system is further reduced.

[0099] In order solve the above-mentioned problem associated with the arrangement of FIG. 5A, the present invention may be applied to a memory system having a serial link bus as shown by way of example in FIG. 5B.

[0100] FIG. 5B illustrates a case where the data bus width of the memory modules 571, 573, 577, and 579 is half of the entire system data bus width. When there are four memory modules, in the conventional semiconductor memory system 500 shown in FIG. 5A, three data buffering operations are needed in transmitting data from the memory module 520 to the memory module 550.

[0101] However, in the semiconductor memory system 560 according to the third embodiment of the present invention shown in FIG. 5B, only one intermediate data buffering operation is needed, such that time delay resulting from data

buffering is reduced greatly, latency of the semiconductor memory system is reduced and the operating speed of the semiconductor memory system is increased. The structure and operation of the memory modules 571, 573, 577, and 579 of FIG. 5B are the same as those of the memory modules according to the first embodiment, and thus, detailed descriptions thereof will be omitted.

5 [00102] As described above, in the semiconductor memory system according to the present invention, the number of memory modules connected to a system data bus is reduced, such that high-speed operation is achieved while large-capacity memory is maintained.

10 [00103] While this invention has been particularly shown and described with reference to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.